DOCUMENT RESUME

ED 157 779 SE 024 781

AUTHOR Helgeson, Stanley L., Ed.; Blosser, Patricia E., Ed.
TITLE Investigations in Science Education, Vol. 3, No. 2. Expanded Abstracts and Critical Analyses of Recent Research.
INSTITUTION Ohio State Univ., Columbus. Center for Science and Mathematics Education.
PUB DATE 77
NOTE 68p.; Contains occasional light and broken type
AVAILABLE FROM Information Reference Center (ERIC/IRC), The Ohio State Univ., 1200 Chambers Rd., 3rd Floor, Columbus, OH 43212 (Subscription, $6.00, $1.75 ea.)

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS *Abstracts; *Critical Thinking; *Educational Research; Evaluation; Instruction; Instructional Materials; Methods; *Observation; Research Methodology; *Science Education; Teacher Attitudes

ABSTRACT
This issue provides analytical abstracts, prepared by science educators, of research reports. The analyses are grouped into three clusters plus three additional analyses. The first cluster, Methods and Materials, contains three studies. The second cluster, Development of Observational Skills in Children, contains three studies. The third cluster, Science Processes, contains two studies. Individual studies focus on formative evaluation, concept learning, and the effects of a teaching experience on the attitudes of preservice teachers. Each abstract includes research purpose, rationale, research design and procedures, findings, interpretations, abstractor's analysis of the research and references. (Author/HM)

INVESTIGATIONS IN SCIENCE EDUCATION

National Association for Research in Science Teaching

ERIC Clearinghouse for Science, Mathematics,
and Environmental Education

Published Quarterly by

The Center for Science and Mathematics Education
College of Education
The Ohio State University
1945 North High Street
Columbus, Ohio 43210

Subscription Price: $6.00 per year. Single Copy Price: $1.75.
Add 25¢ for Canadian mailings and 50¢ for foreign mailings.

TABLE OF CONTENTS

Volume 3, Number 2

TABLE OF CONTENTS (Continued)

NOTES . . .

from the Editor

Analyses of research reports are grouped into three clusters in this issue of ISE, plus three additional analyses. The first cluster, METHODS AND MATERIALS, contains three studies. The second cluster, DEVELOPMENT OF OBSERVATIONAL SKILLS IN CHILDREN, contains three studies conducted by the same pair of investigators and analyzed by one reviewer. The third cluster, SCIENCE PROCESSES, contains two studies. Individual studies focus on formative evaluation, concept learning, and the effects of a teaching experience on the attitudes of preservice teachers.

Publishable responses to the analyses and to the grouping of studies are encouraged.

Stanley L. Helgeson
Editor

Patricia E. Blosser
Associate Editor

METHODS AND MATERIALS

Wideen, Marvin. "Comparison of Student Outcomes for Science--A Process Approach and Traditional Science Teaching for Third, Fourth, Fifth and Sixth Grade Classes: A Product Evaluation." Journal of Research in Science Teaching, 12(1):31-39, 1975.

Descriptors--*Academic Achievement; *Curriculum Evaluation; *Elementary School Science; Elementary Education; Process Education; *Program Evaluation; Science Education; Science Course Improvement Project; *Student Attitudes

Expanded Abstract and Analysis Prepared Especially for I.S.E. by David P. Butts, The University of Georgia.

Purpose

The purpose of this study was to document the effectiveness of Science--A Process Approach in contrast with "traditional science" teaching. Effectiveness was defined as including outcome variables of knowledge, process skills, interests, attitudes, and students' view of the classroom. In addition to the curriculum, student characteristics of intelligence, gender, and grade level were also thought to be related to performance on the outcome variables.

Rationale

Results of previous research suggest that outcome measure of science instruction have not been consistently found to be related to the curriculum. The implied contextual model is that summative evaluation should be done to show how student knowledge and attitude outcomes are related to the content of instruction.

Research Design and Procedure

Using a non-equivalent control group design, 531 students from grades 3 to 5 in two rural school districts were pre- and posttested on six measures related to achievement and attitude. The experimental classes had instruction based on Science--A Process Approach for an unspecified time per week and for an undesignated number of weeks. The control group

received "traditional instruction" for a similar unspecified time per week for an unknown duration. Reliability of instruments are reported. No estimates of the validity of the measures for the variables of this study are given.

Findings

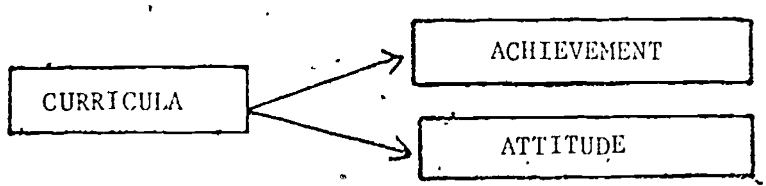
Students in rural South Dakota (third to sixth grade) who had the experimental treatment showed a greater growth in science knowledge. No differences in interest in science were found related to achievement and no differences in the students' perception of the classroom were found when compared with similar groups in "traditional" classes.

Interpretation

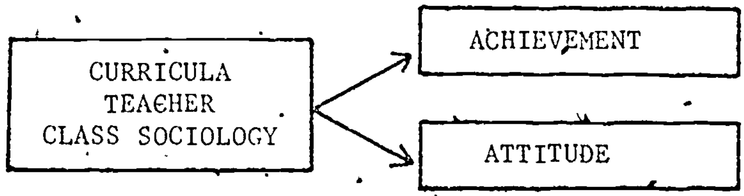
Interpretations are focused on discussing how similarity in performance on the outcome measures may be indicative of the fact that they are measuring different variables than those as originally expected. Uncontrolled factors in the school context are also suggested as possibly influencing the results.

ABSTRACTOR'S ANALYSIS

In his introduction, the author identifies that previous studies have examined several variables which may be predictors of achievement: socio-economic background, age, curriculum, reading readiness, gender. He also noted that specific student variables of creativity, interest, and reading readiness may also be influencing and influenced by the students' learning. What is missing from this study is an explicit framework on which to hang the previous research or the questions of this study. The inferred framework is



By the end of the study, recognition is given to the fact that "curriculum" is likely too global a variable and should include at least a recognition that the school, the teacher, and the class sociology are key factors in predicting either achievement or attitude.



In a valid study, one has confidence that the relationships described in the conclusions are indeed present. In this study, experimental and control groups are identified. The extent to which these were different or the validity of their difference is omitted. In the absence of information about <u>what</u> specifically was done to <u>whom</u>, <u>when</u>, the limitations with which the conclusions can be believed are enormous. If this study is intended to provide the reader with a summative view of how program A compares with program B, the very globalness of the treatment leaves the results so tentative that they are extremely limited in their usefulness. An alternative and stronger design is one that would have documented specific changes in achievement for identified objectives after an explicitly described treatment.

In the absence of explicit objectives for either treatment or control, why one would expect differences in the range of outcome variables is omitted from the manuscript. However, the clarity of the written report is refreshing.

In studies where instruction is being evaluated, the changes instruction is expected to produce under explicitly described circumstances and with what kinds of students must be included. The omission of each of these three descriptors seems to be amply illustrated in the literature. Further studies would provide much more meaningful data and be more useful to decision-makers if they

1) described the specific outcomes of the instructional program which are

2) to be achieved in smaller time periods than "an entire year" and

3) under replicable conditions with three identified variables known to influence the impact of instruction; e.g., the teacher, the context, and the student.

Nelson, Bess J. and Arthur L. White. "Development of a Path-Analysis Model Relating Elementary Teacher Variables and Science Teaching Practices." Journal of Research in Science Teaching, 12(4): 379-384, 1975.
Descriptors--Educational Research; Elementary Education; *Elementary School Science; *Instructional Innovation; *Instruction; Science Education; *Statistical Surveys; *Statistical Analysis; Teaching Methods

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Frances Lawrenz.

Purpose

The purpose of this article was to examine the relationship between elementary science teacher variables and science teaching practices. The investigation proceeded in three steps. First the authors identified the relationships, then they hypothesized a model to describe and explain them and, finally, they tested the model.

The data employed in the study were from the elementary teacher segment of the data from a national survey of science teaching practices conducted by the Faculty of Science and Mathematics Education in cooperation with the Science Mathematics and Environmental Education ERIC Clearinghouse at the Ohio State University.

Rationale

The article provides no discussion of the theoretical basis for the investigation although the underlying assumptions must have been that there were indeed relationships between science teacher variables and teaching practices and that a path model analysis of teacher questionnaire data was a good vehicle for examining these relationships. Also implicit in the study was the hope that the formulation of a model of the relationships would have some value, i.e., that it would contribute to a more accurate conceptualization of the interaction between teachers and teaching.

Research Design and Procedure

A sample of 2,948 schools was drawn from a population consisting of all the public elementary schools in the Mideast, New England, and Southwest regions of the United States during the 1970-71 school year. From these, 1,444 schools responded and usable data were received from 880. Because initial analyses indicated no difference between the three sampled regions, teacher questionnaire data from one region, the Mideast, were selected for this investigation. Through factor analysis, the 72 teacher questionnaire items that discriminated among respondents were reduced to eight factors: Factor 1, the Availability of Supply and Equipment Budget; Factor 2, Teaching Experience; Factor 3, Use of Science Course Improvement Projects; Factor 4, Time Allotment for Science; Factor 5, Use of Alternatives to Standard Instructional Programs; Factor 6, Use of Audio-visual Aids; Factor 7, Science Teaching Knowledge; Factor 8, Science Inservice Participation. These eight factors were then used as the variables for forming a path analysis model.

A total of eight models were constructed. The first was built using the intra factor correlations from the factor analysis for the Mideast region. This initial model graphically presented the eight factors with lines joining those with correlations significant at $p \leq .05$. Causal direction was hypothesized and arrows indicating this were included in the diagram. For example, it was hypothesized that Factor 1 (the availability of supplies and equipment budget) caused Factor 8 (Science Inservice Participation) so an arrow was drawn from Factor 1 to Factor 8.

Once the causal direction of the relationship between factors had been hypothesized, the formulation of a path analyses model using path regression coefficients was possible. These path regression coefficients were determined following Blalock's equation and represented the effects of one variable on another with the effects of the other variables held constant. The first path model was drawn using regression coefficients obtained from the Mideast data. The eight factors were presented in a diagram with directed lines connecting those with significant ($p \leq .05$)

regression coefficients. Two more models were then constructed; one using the data for the New England region and one using the data for the Southwest region. Finally, to test the reliability of the method of model development, four more models were constructed by dividing the data in the Mideast region in half by two methods and forming a model for each. The two methods of splitting were odd and even numbered cases and every other pair of cases.

Findings

Because each of the seven path analysis models showed somewhat different relationships among the factors, it was concluded that the relationships between the eight factors vary for the given population and that the models fluctuate with the sample. The source of these differences was explained as being due to sampling or regional differences or both. However, these results are confounded and could not be separated. In spite of the differences among the models, they all showed a relationship between Availability of Supply and Equipment Budget (Factor 1) and Science Inservice Participation (Factor 8). In three of the seven path models, Factor 1 was connected to Factor 6 (Use of AV Aids) through Factor 8 and in four of the models Factor 1 was connected to Factor 3 (Use of science Course Improvement Projects) through Factor 8.

Interpretations

The authors concluded that the availability of a budget for supplies and equipment may have a directional effect through science inservice participation on the use of science course improvement projects, audiovisual aids and/or alternatives to standard instructional programs. In other words, an adequate budget may not be sufficient motivation to get teachers to use innovative teaching techniques, inservice training may also be necessary.

Three implications were suggested from the results of this study.

1. Elementary science teachers need to have budgeted amounts for supplies and equipment.

2. Inservice and preservice teacher-education programs available to elementary science teachers should acquaint them with the materials and help them develop the skills necessary for good science teaching, i.e.:

   a. Use of Audiovisual Aids (Factor 6)
   b. Use of Alternative to Standard Instructional Programs (Factor 5)
   c. Use of Science Course-Improvement Projects (Factor 3).

3. Inservice and preservice teacher-education programs should provide guidance as to how teachers can influence budgeting for supplies and equipment in their school system.

ABSTRACTOR'S ANALYSIS

The path analysis model utilized by Nelson and White is an interesting way of examining the traditional question of the relationship between teacher variables and teaching practices. It offers some unique possibilities. Since path analysis is an uncommon technique in science education research, it may have been useful if the authors had included a discussion of the methodology involved and its unique advantages for the investigation of the teacher/teaching relationship. Generally, path analysis is a method for producing a schematic diagram of related variables. Formulation of a model provides a specific link between an *a priori theoretical notion of causal* connections and quantified estimates of causal impact. The method presents the same type of information provided by multiple regression or correlation techniques but with a causal component. The advantages of path analysis are its graphic nature of presentation, the opportunity to hypothesize causation, and the demonstration of mediating variables. Its limitations are the less than straightforward nature of the results and the problem of hypothesizing causal direction. This problem is particularly acute when no time constraints or empirically determined information is available to concretely identify causation. As in this study, for example, it may be just as

likely that teacher Science Inservice Participation (Factor 8) would cause an increase in the Availability of the Supply and Equipment Budget (Factor 1) as the other way around.

While the research design is sound, there are a few instances where more information would have been helpful. No doubt these gaps could be filled in by reading Nelson's thesis but it would be more convenient if they had been included in the article. The authors do not report how the sample was drawn, how the data were collected or, since the data were analyzed by region, how many respondents were in each region. There is no mention of a nonrespondent study nor of what might have made over half of the collected data unusable. Also, the comparative analyses between the regions are not described nor is the method of factor analysis.

The interpretation of the results is somewhat confusing. Perhaps this is due to the lack of any discussion leading to the formulation of the research plan. It would have been beneficial to see how the present study fit into the general context of investigative work on teacher variables. Such an in-depth analysis would also have helped to support the causal directions applied to the model. In this same vein it would have been instructive if the authors had devoted some time to a discussion of the conditions in the elementary science classroom. More attention to other research or a theoretical basis would also have helped to flesh out the conclusions and implications sections of the article.

It was difficult to understand why eight models were necessary. If the regions were known to be the same, why should a model be formed for each? Why were two sets of halves formed? Possibly the generation of several models was an attempt to empirically verify the initial model. Usually, however, in path analysis a theoretical model is proposed and then data are obtained to test the validity of the model; the theoretical relationships are quantified. In this study the authors seem to pay little attention to the extent of the relationships indicated by the models and concentrate instead on the number of times the relationship appeared.

In summary, it seems that the article provides important and relevant data but that it is not presented in the context of previous research. Also, although the method of analysis employed in the study offers interesting possibilities for future research, it is not described in detail.

Daugs, D. R. "Influence of Multilevel Science Materials on Achievement of Sixth Grade Students." Journal of Research in Science Teaching, 10(2):147-152, 1973.

Descriptors--*Academic Achievement; *Earth Science; Educational Research; Elementary School Science; *Instructional Materials; *Media Selection; Reading Difficulty; *Reading Level; Science Education

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Eugene Chiappetta, University of Houston.

Purpose

The study was designed to investigate the hypothesis that students will make greater gains in achievement if they receive instructional materials at their reading level than if they receive instructional materials that are above their reading level. To elucidate this hypothesis the following questions were researched: (1) Is there a significant difference in achievement on criterion test items between classes of sixth grade students who study material written at one level of difficulty and classes of sixth graders who study material written at five levels of difficulty when students are placed at their respective reading levels by an informal reading inventory?, and (2) ~~Is there~~ a significant difference in achievement on criterion test items between classes of sixth grade students who study material written at one level of difficulty and classes of sixth grade students who study materials written at five reading levels when students are placed at their respective reading levels by an informal reading inventory when classrooms are statistically equated with respect to teacher rating?

Rationale

Daugs was investigating the importance of matching the science test materials to the reading level of the intended student. He analyzed Fryback's work which dealt with the use of multilevel reading materials by placing fifth grade students at one of five reading levels in SRA's "The Earth's Atmosphere." Fryback reported no significant difference in gain between

classrooms using five levels or using one level of material. At first Daugs questioned the use of the Standardized Metropolitan Reading Test as a means of placing students at their appropriate reading levels. However, after investigating the results of assigning students to their respective reading levels using an informal reading inventory, Daugs hypothesized that the teacher variable was masking the effects of matching text material to student reading level. Daugs reasoned that in the classrooms where all students were exposed to the same reading level materials, the teacher assisted students who were disadvantaged by the materials above their level.

Research Design and Procedure

Fifteen classrooms of sixth grade students participated in the study. Eight classrooms were assigned to the experimental treatment using multi-level materials and seven classrooms were assigned to a control group using one level of material. A total of 368 pupils participated in the study. The pupils were assigned to reading levels using an informal reading level inventory based on the science materials actually used in their classrooms (SRA's "The Earth's Atmosphere"). The Metropolitan Reading Test was used to equate the two treatment groups on general reading ability.

The experimental design was a pretest-posttest control group design described by Campbell and Stanley. The design is designated:

$$\begin{array}{cccc} R_1 & O_1 & X & O_2 \\ R_2 & O_3 & & O_4 \end{array}$$

Where R_1 and X represent the eight sixth grade classrooms using the SRA science reading materials written at five levels of difficulty.

R_2 represents the seven sixth grade classrooms using science reading materials written at one reading level. The Metropolitan Reading Test

was used as a measure of general reading ability and the scores from this test were used to equate the two treatment groups.

The Cloze readability procedure was employed as the criterion test for the pretest and posttest measures. The Cloze procedure is generally used to measure reading comprehension and can be used to assess subject matter achievement, as was the case in the present study. With the Cloze procedure every fifth or seventh word is deleted from textual material and the student is expected to fill in the missing word, thus demonstrating his knowledge of the material.

The teachers' ability to manage reading instruction was believed to be a variable contributing to the effectiveness of the reading program. For this reason, the teachers were rated on criteria in 13 subcategories as to their effectiveness in managing the treatment reading programs. The ratings were made by an observer visiting the classrooms during the science reading instruction. The mean teacher rating scores were used to equate the classrooms statistically by means of the analysis of covariance procedure, with teacher rating as the covariate.

Findings

Analysis of variance procedures revealed no significant difference between experimental and control groups on either the subtest gains or on total gain scores. There was a significant difference between teacher management ratings of the experimental and the control groups. The teachers in the experimental treatment who managed five levels of reading material received the lowest ratings.

Interpretations

The results of this study indicate that matching pupils' reading levels by employing five levels of science textual materials is no more effective than using one reading level. This result confirms earlier research

findings. Another result of the study was that the teachers who employed the multilevel reading program were rated lower on their ability to manage this instructional procedure than the teachers who employed a single level reading program. However, the apparent variation in teacher effectiveness had no effect on the treatment.

ABSTRACTOR'S ANALYSIS

Upon reviewing the present study one is taken back by the fact that the multilevel approach to science instruction was no more effective than the single reading level approach. Initially one would expect the research to substantiate the utility of the multilevel approach. This expectation is built upon several factors. First, on intuitive grounds the multilevel approach appears superior to the single level approach. Second, the present study is a refinement and an extention of a previous study performed by Daugs and other researchers cited by Daugs. Third, the experimental pretest-posttest research design is a good design and employed well by Daugs. Given the above, why was there not a significant difference in the treatments?

There are at least two possible reasons for treatments producing similar results. The first has to do with the duration of the treatments. The research report does not indicate how long the students participated in the reading instruction. If the reading instruction was relatively short in duration, it is conceivable that the enhancing effects of the multilevel approach would not be demonstrated. If, however, the reading instruction for both treatments was long in duration, say several months for several hours each week, then the advantage of the multilevel treatment might be evidenced.

Another reason why the multilevel treatment did not live up to expectation might be that many other non-reading activities were contributing to the achievement displayed by both groups. Presumably there were laboratory activities associated with the SRA science reading lessons which benefitted both groups. There were probably classroom discussions

related to the science activities that pupils in both groups participated in. In addition, there occur numerous activities in school, inside or outside of language arts classes, which pertain to language arts that would influence the type of achievement measured by the Cloze procedure. Hence, if other activities were dominant to the science reading experiences, the effects of the multilevel approach would be masked.

The present line of research should be continued. It is important to the profession to better understand the relationship between reading materials and science achievement. Subsequent studies should employ a reading treatment that is long in duration and that is dominant to other activities that might impart similar knowledge. In this manner the true effects of the reading treatment(s) can be ascertained.

DEVELOPMENT

OF

OBSERVATIONAL SKILLS IN CHILDREN

Barufaldi, James P., and Maureen A. Dietz. "The Performance of Children on Visual Observation and Comparison Tasks." Science Education, 59 (2):199-205, 1975.

Descriptors--*Achievement; *Educational Research; Elementary Education; *Elementary School Students; Learning Activities; Science Education; *Visual Learning; *Visual Perception; Visual Stimuli

Expanded Abstract and Analysis Prepared Especially for I.S.E. by William S. La Shier, Jr., University of Kansas.

Purpose

The following questions were investigated in this study:

1. Does the use of different types of visual stimuli (solid objects, photographs of the objects, and drawings of the objects) result in different observation and comparison task scores among groups representing the same or different grade levels?
2. Are there differences between the mean scores of boys and girls on each of the four observation and comparison task areas of color, size, form and form-detail?

Rationale

The development of observation and comparison skills was cited as a major objective of many of the elementary science programs. These skills require that students collect information about the physical attributes of an object and recognize similarities and differences among these attributes chiefly through the sense of sight.

Research by Stevenson and McBee (1) indicated that young children discriminated size more efficiently when using three-dimensional objects than when presented with pattern or slightly raised objects. Dornbush and Winnick (2) found that young children learned to discriminate between two solids, cube and parallelopiped, faster than between a square and rectangle. These and related studies suggested that the performance of children on tasks of observing and comparing when presented with a solid object may differ from

that demonstrated when they are presented with a two-dimensional representation of the object. Thus, the results of this study may provide the science educator with direction in the choice and nature of instructional materials used to teach process skills.

Research Design and Procedure

Three treatment groups of elementary school students were formed to respond to a set of 14 visual observation and comparison tasks which were administered orally. Two solid objects, a cube and a cylinder, served as a visual stimulus for the first treatment group as it responded to the 14 tasks. The second treatment group had photographs of the cube and a cylinder to refer to in responding to the questions. The third treatment group had drawings of the objects as their frame of reference in responding to the 14 tasks. Three tasks focused upon the physical attribute of color, two dealt with size, five tasks focused upon form and four tasks dealt with form-detail. A score of 1 was assigned to each correct response to a task and a zero to incorrect responses. A total score and four subscores were calculated.

Two hundred forty randomly selected students from grades one, two, four, and six were used in this study. The subjects were from two urban elementary schools with a racial composition of 99 percent Black and 1 percent Oriental and others. The overall measure of task reliability, Kuder-Richardson No. 20, yielded a value of 0.664.

The posttest only research design called for a series of two-way analyses of variance with three types of visual stimuli as treatments and four grade levels. A separate two-way analysis of variance was performed for each of the four dependent variables of color tasks, size tasks, form tasks, and form-detail tasks. Post-hoc comparisons were performed to locate specific sources of significance indicated by the two-way analyses of variance.

Separate one-way analyses of variance were used to determine if there were any significant differences between male and female responses to each of the four sets of tasks.

Findings

The results of the study indicated that there were no differences among the three treatment groups in terms of the mean number of correct resonses to the observation and comparison tasks dealing with color. There were, however, significant differences between first graders and sixth graders to the tasks employing color, with the older children giving more correct answers.

The three treatment groups (solid objects, photographs, drawings) did not differ significantly on the tasks employing the attribute of size. As with the color tasks, a significant difference was noted between age levels with older children giving more correct answers.

The three treatment groups differed significantly on the tasks employing the attribute of form. Post-hoc comparisons of treatment effects indicated that the use of solid objects or photographs resulted in more correct student responses than the same tasks employing drawings. Again, significant differences were noted between age levels with six graders making more correct responses on the form tasks than the first, second, and fourth graders.

The three treatment groups differed significantly on the tasks employing form-detail. Children made more correct responses on the form-detail task when viewing the photographs in contrast to viewing the drawing of the objects. Significant differences were noted among the grade levels with sixth graders making more correct responses than first, second, and fourth graders.

When the mean number of correct answers of males and females were contrasted for each of the four sets of tasks, only one significant difference was noted; females made more correct responses on the color tasks than did the males.

Interpretations

The mean scores were higher on tasks which focused on the attribute of form when children observed and compared either the solid objects or the photographs of the objects, in comparison to children using drawings as the stimuli. Post-hoc comparisons of treatment effects indicated that children made more correct responses on the form-detail tasks when viewing the photographs than when they viewed the drawings of the objects. Mean scores on color or size tasks were not significantly different among the three treatment groups.

The older the children, the more successful they were in performing the tasks. The females surpassed the males on the tasks related to color. The implication is that males in the elementary school may have to be provided with additional experiences with color observation and comparison tasks.

The implication of the study focuses on the need for careful selection of appropriate teaching material and method with consideration given to the age and sex of the learner as well as the color, size, form and form detail among the instructional materials.

ABSTRACTOR'S ANALYSIS

This study is related to two other studies, Dietz and Barufaldi (1975) and Barufaldi and Dietz (1975b), which are both critiqued in the present ISE volume. In the first study, Dietz and Barufaldi (1975) studied the effect of presenting novel vs. ordinary objects on student observation skill. The observation skills of youngsters were also tested under the conditions of comparing two objects vs. presenting one object at a time. The final activity required students to recognize both similarities and differences of objects with the objective of determining whether similarities were mentioned more than were differences.

The Barufaldi and Dietz (1975) study investigated the differential effect of using either solid objects or two-dimensional representations of objects.

on the subsequent performance of youngsters on visual observation and comparison tasks. These two previous studies both used basically the same sample population of youngsters selected from one elementary school.

The present study concisely outlined the effects of different types of visual stimuli on the performance of children on visual observation and comparison tasks. The article would have been clearer, however, if a table of mean scores by grade level had been included.

The authors spoke of the educational implications of the study but provided only a few concrete suggestions for teachers such as the need to provide males in the elementary grades with additional experience with color observation and comparison tasks.

In a recent descriptive article, Thier (1976) described alternative approaches for developing visual perception skills through the use of selected science experiences. Marlene Thier, an elementary school learning disability teacher, provided some suggestions to enable a student to overcome visual perception problems through involvement in science activities in many of the new curricula. She also pointed out that the skills needed to work with SCIS Mr. O, for example, are related to reading. To read successfully, a child must be able to discriminate between words that are similar and words that are different, along with other related skills.

Another instructional approach would be to directly teach for visual perception. Ritz and Raven (1970) investigated the impact of science instruction and direct instruction in visual perception on the subsequent reading readiness of kindergarten children. In their study, three groups of kindergarten students were used to investigate whether AAAS Science - A Process Approach or instruction in a portion of perceptual training program would enhance the reading readiness attainment of the students. The third group of youngsters received instruction in both the AAAS and the Frostig material. None of the three treatments produced significant differences and the researchers concluded that science and/or visual perceptual training can be included in kindergarten programs without impairing the reading readiness of the youngsters.

The reluctance of Barufaldi and Dietz, in the present study, to generalize the findings to other groups was justified because the sample population was composed of 99 percent Black students from a neighborhood in which 56 percent of the contacted families reported incomes of less than $5,000 per year. The socioeconomic status of the students may have influenced their performance on the various visual tasks. Lowery and Allen (1970) examined the performance of upper, middle, and lower socioeconomic groups of first graders on 35 tasks of resemblance sorting. This term meant grouping together two objects because they are similar in some way. The nonverbal test booklet consisted of drawings and figures. The findings indicated that low socioeconomic group scored significantly lower on the sorting tasks than the other two groups. Also, a hierarchial sequence of difficulty for the sorting task was evidenced.

REFERENCES

Barufaldi, J. P. and Maureen A. Dietz. "Effects of Solid Objects and Two-Dimensional Representations of the Objects on Visual Observation and Comparison Among Urban Children." Journal of Research in Science Teaching, 12(2):127-132, 1975b.

Dietz, Maureen A. and J. P. Barufaldi. "Observations and Comparison Tasks Using Ordinary and Novel Objects." Journal of Research in Science Teaching, 12(2):193-198, 1975.

Dornbush, R. L. and W. A. Winnick. "The Relative Effectiveness of Stereometric and Pattern Stimuli in Discrimination Learning in Children." Psychonomic Science, 5:301-302, 1966.

Lowery, Lawrence F. and Leslie R. Allen. "Socio-Economic Status and Sex Differences in Visual Resemblance Sorting Tasks at the First Grade Level." Journal of Research in Science Teaching, 7(1):57-65, 1970.

Ritz, William C. and Ronald Raven. "Some Effects of Structured Science and Visual Perception Instruction Among Kindergarten Children." Journal of Research in Science Teaching, 7(3):179-186, 1970.

Stevenson, H. W. and G. McBee. "The Learning of Object and Pattern Discrimination by Children." Journal of Comparative Psysiological Psychology, 51:752-754, 1958.

Thier, Marlene. "Utilizing Science Experiences for Developing Visual Perception Skills." Science and Children, 13(6):39-40, March, 1976.

Dietz, Maureen A., and James P. Barufaldi. "Observations and Comparison Tasks Using Ordinary and Novel Objects." Journal of Research in Science Teaching, 12(2):193-98, 1975.
Descriptors--Educational Research; Elementary Education; *Elementary School Science; *Observation; Science Education; *Task Performance; *Visual Perception; *Visual Stimuli

Expanded Abstract and Analysis Prepared Especially for I.S.E. by William S. LaShier, Jr., University of Kansas.

Purpose

The problems investigated in this study are related to the following three questions:

1. Do children demonstrate greater skill in observing an ordinary object than in observing a novel object?
2. Do children demonstrate greater skill in observing two objects, presented one at a time, than in comparing the two objects?
3. Do children demonstrate greater skill in recognizing differences between objects than in recognizing similarities in objects?

Research Design and Procedures

This posttest only research design was used with 66 randomly selected students in grades one through six from one school in an eastern city. The students apparently were from low income families. Eleven first grade students responded to four tasks presented in an interview situation. This same procedure was followed using an equal number of students at grades two through six. The interview data were recorded by a cassette tape recorder and later analyzed.

In Task 1, the student was handed either a cube or a ball and asked questions to elicit observational statements. In Task 2, either a gyroscope or kaleidoscope was handed to the student and observational

statements were again asked for. In Task 3, the student was handed the two objects previously presented and was asked, "How are they alike?" The final task was to ask the student, "How are these two objects different?" The total number of correct responses were determined for each grade level using a procedure which had been modified from that of The Inquiry Skills Measures (Nelson, 1971).

Findings

The students were equally skillful in making observations of an object regardless of whether it was familiar or novel. The results also indicated that children demonstrated significantly greater skill in observing than in comparing. There was also a significant difference among grade levels in the performance of children on observation and comparison tasks.

When the student responses to the comparison tasks were distinguished as either being a similarity or a difference response, a significant difference was noted which favored more student statements based on differences than those based on similarities. No significant differences were noted among the grade levels.

Interpretation

In helping students to develop observational skills the selection of either a novel object or a familiar object did not seem to be crucial at any grade level. The study also revealed that children at all grade levels made at least twice as many observations as comparisons. The researchers suggested that perhaps comparison skills tasks should be stressed more during instruction. When the student responses related to either similarities or differences, the children used differences more often. This suggests that children should be given further training in recognizing similarities in objects, particularly since similarities are the basis for developing classifying skills.

ABSTRACTOR'S ANALYSIS

The design of this study was explained well and the only limiting feature was that the sample population seemed quite narrow in terms of socio-economic status and geographic distribution. The results of Task 3 were quite intriguing. Why do subjects make more statements based on differences than on similarities? Could the sequence in which the questions were posed make a difference?

The present study indicated that the introduction of novel objects did not result in the students making more observations. In a study by Gillespie (1970) the use of ordinary objects was recommended.

In Gillespie's study a group of 96 youngsters, divided by age (five, six, seven, and eight years) and sex into equal groups, were administered tests of discrimination between leaves, association of leaves with pictures of leaves, sorting leaves into generic groups, and communication of concepts of leaf structure. No sex difference or interactions were found. However, there was a significant age effect on all tests, with older children having less difficulty than did the five year olds. It was recommended that children be taught to recognize the more common species of trees through first-hand experience with leaf material, starting from age eight.

In another study, the concept of student curiosity about science objects was examined. Peterson and Lowery (1972) unobtrusively assessed the amount of coordinated sensory-motor activity of 120 school children (ages 5 through 13) directed toward an array of science objects in a waiting room. The researchers were interested in the amount of curiosity expressed through motor activity. This curiosity was found to be significantly related to groups based on racial-ethnic origins, but not to groups based on age or sex. Black children expressed significantly more curiosity while waiting than did non-Black children. A significant interaction effect of age, sex, and racial-ethnic origins was also found. Interaction effect was due to a rapid increase in curiosity from kindergarden to sixth grade among Black males and a gradual decrease in curiosity among non-Black males over the same time period.

Peterson and Lowery also reported that children who exhibited greater amounts of curiosity usually asked few unsolicited questions. One wonders if in the Dietz and Barufaldi study, whether the non-significant differences in number of student observation statements when presented with either novel or ordinary objects might be a result of the constraints associated with the verbal question format. Perhaps a follow-up study might deal with relative amount of curiosity associated with students observing first a novel object and then a familiar object.

REFERENCES

Gillespie, James Pittard. "A Study of Concept Development and Communication Abilities in Children Aged Five Through Eight Using Leaf Morphology as the Vehicle." Doctoral dissertation, George Peabody College for Teachers, 1970.

Nelson, M. A. and E. C. Abraham. "The Inquiry Skills Measure." Paper presented at the annual convention of the National Association for Research in Science Teaching, Silver Springs, Maryland, March, 1971.

Peterson, Rita W. and Lawrence F. Lowery. "The Use of Motor Activity as an Index of Curiosity in Children." Journal of Research in Science Teaching, 9(3):193-200, 1972.

Barufaldi, Jam[redacted]P., and Maureen A. Dietz. "Effects of Solid Objects and Two-Dimensional Representations of the Objects on Visual Observation and Comparison Among Urban Children." Journal of Research in Science Teaching, 12(2):127-132, 1975(b).

Descriptors--Developmental Tasks; Educational Research; Elementary Education; *Elementary School Science; *Models; *Observational Learning; Science Education; *Urban Youth; *Visual Perception

Expanded Abstract and Analysis Prepared Especially for I.S.E. by William S. LaShier, Jr., The University of Kansas.

Purpose

The problems investigated in the study were related to the following two questions:

1. Do the differences in visual perception of solid objects and two-dimensional representations of the objects affect the performance of children on visual observation and comparison tasks?

2. Do the different types of two-dimensions representations (photographs or drawings) affect the performance of children on visual observation and comparison tasks?

Research Design and Procedures

The posttest only research design contrasted the responses of three groups of students; each group observed and compared one type of visual stimulus (solid, photograph, or drawing). That is, one group of students would work with either a set consisting of a cone and pyramid or the set composed of a cube and cylinder. The second group of students would work with photographs of one of the sets of objects. The third group worked with drawings of the sets of objects. Each of the three groups at each of four grade levels, 1, 2, 4 and 5, consisted of 17 to 20 students. The procedure in testing a particular group was to place one of the objects or representations on the desk of each subject and ask the student to respond to ten observation tasks. The second object or representation was then introduced and the student responded to an additional 16 tasks

requiring the comparison of two objects or representations. A score of one was assigned to each correct answer with the individual's total score being the total number of correct responses. The 228 children selected for the study were from one elementary school in a large eastern city and were randomly selected from grades one, two, four and six.

Findings

Comparison of group means through the computation of Bonferroni t-values indicated that in grade two the students who observed and compared the solid objects scored significantly lower than did students using the photographs. In grade four, the groups using either the solid objects or photographs did significantly better than did the group using the drawings. In grade six, the group using the solid objects did significantly better than did either the group using the photographs or the group using the drawings.

Interpretation

Children in grades four and six were more skillful on observation and comparison tasks employing solid objects than on those tasks using drawings of the objects. In addition, fourth graders were more successful when using photographs than when presented with drawings.

The authors suggested that perhaps educators should design experiences that would give children greater opportunity to utilize photographs and drawings in the development of scientific skills.

ABSTRACTOR'S ANALYSIS

A review of research related to the three articles reviewed in this series began with an on-line computer search of the ERIC system from 1966 through December 1977 using key identifiers listed in the Thesaurus of ERIC Descriptors (1977). The terms, elementary school science and elementary school mathematics, were combined and these represented 5,008 entries

labelled as Set 3. The intersection of Set 3 with the set of entries related to the term visual stimuli produced five entries for Set 5. The sets relating to visual learning but not Set 5, and the set related to visual perception but not Set 5 produced 8 articles and 15 articles, respectively, when intersected with Set 3. Two of the articles found in this search suggest research directions somewhat related to the reviewed article.

Champagne (1970) reported a high positive correlation between the ability of kindergarten students to draw plane geometric figures and their ability to conserve mass in the Piagetian sense. The study was based on the proposition that the child's attention to misleading visual clues is an important factor in his lack of ability to conserve mass and that instruction in drawing two-dimensional figures should affect ability to conserve mass.

Walker (1972) reported that no significant differences were observed in problem solving behavior among three groups of sixth grade students. One group received visual automated instruction, the second group received automated auditory instruction and the third group received no instruction. This study perhaps suggested an area of research not varied in the three Baryfaldi and Dietz studies. That is, perhaps the manner in which the tasks are presented could be varied. Another possible dependent variable in these studies could be time spent on each task, particularly if the instructions were automated for the different grade levels.

REFERENCES

Champagne, Audrey Ann Briggs. "An Investigation of the Effectiveness of Visual-Motor Experiences in the Development of the Ability to Conserve Mass." Doctoral dissertation, University of Pittsburgh, 1970.

Thesaurus of ERIC Descriptors, 7th Edition. New York: Macmillan Information, Macmillan Publishing Co., Inc., 1977.

Walker, Mary Ann. "The Comparative Effects on Two Methods of Automated Instruction, One Visual and One Auditory, in Teaching Selected Problem Solving Behaviors to Two Groups of Sixth Grade Students." Doctoral dissertation, St. Louis University, 1972.

SCIENCE PROCESSES

Seymour, Lowell A. and Frank X. Sutman. "Critical Thinking Ability, Open-Mindedness, and Knowledge of the Processes of Science of Chemistry and Non-Chemistry Students." Journal for Research in Science Teaching, 10(2):159-163, 1973.

Descriptors--*Chemistry, *Critical Thinking, *Curriculum Evaluation, *Educational Objectives, Educational Research, Instructional Materials, Science Education, Secondary School Science, *Student Attitudes

Expanded Abstract and Analysis Prepared Especially for I.S?E. by J. Dudley Herron, Purdue University.

Purpose

The purpose of the study is to provide a summative evaluation of a curriculum guide developed by the Chemistry Curriculum Committee of the Philadelphia School District.

"Instruction based on this Guide was intended...to develop critical thinking ability, open-mindedness, and knowledge of the processes of science" (p. 159). The evaluation was in terms of these constructs.

Rationale

Essentially, the rationale for the investigation was that critical thinking ability, open-mindedness, and knowledge of science processes are important goals of sicnnce instruction. This is not a researchable question but the authors do cite several references purported to endorse such goals. The goals are ones likely to be endorsed by a majority of science educators.

Research Design and Procedure

The design used was a nonequivalent control group design (Design 10, Campbell and Stanley (1963). The sample consisted of eleventh grade students in two schools, drawn from a population of 16 schools using the guide.

Data were presented which suggest that the sample was representative of the population on several variables.

Eleventh grade chemistry students in the two sample schools comprised the experimental group. The control group consisted of other eleventh grade students matched to chemistry students by sex and intelligence.

The number of subjects in the study is not mentioned but the degrees of freedom listed for the various F-ratios and t-tests ranged from 142 to 182, suggesting that the number in each group was in the order of 150-200. The authors are not explicit about the duration of the study. One comment suggests that it was one semester.

Three tests were administered to both groups as pretests and posttests. These were:

a) Watson-Glaser Critical Thinking Appraisal From YM (WGCTA)
b) Rokeach Dogmatism Scale - From E (RDS)
c) Wisconsin Inventory of Science Processes (WISP)

Scores from pretests and posttests were used to make two comparisons. First, the pretest and posttest scores for students who used the guide (experimental group) were compared. Second, posttest scores for the experimental and control groups were compared.

Findings

The experimental group made statistically significant gains from pre- to posttest on the WGCTA and the posttest scores of the experimental group were significantly higher than those of the control group. (No means are reported. "Significant" as used here refers only to statistical significance at the 0.05 level of confidence or beyond. One is unable to judge the educational significance of the differences found and the authors make no comment in that regard.)

The change in mean on the RDS from pre- to posttest for the experimental group was not significant at the 0.05 level, but there was a difference in posttest means for the experimental and control group. The difference apparently indicated more open-mindedness on the part of the experimental group. Since means are not reported, this result is somewhat ambiguous. The authors report no significant difference in pretest mean for experimental and control groups, no significant gain by the experimental group, but a significant difference in posttest means for the two groups. Whether this is due to an increase in dogmatism by the control group or small initial differences in favor of the experimental group indicate a decrease in knowledge of the processes of science but the comparison of posttest scores for experimental and control group shows no significant difference as did the comparison of pretest scores.

The authors conclude that the Guide did not result in an increase in knowledge of science processes. This appears to be a safe assumption, in spite of the ambiguous results.

Interpretations

The authors interpret these results to indicate that use of the Guide results in an increase in critical thinking ability but not in an increase in knowledge of science processes. They consider the results pertaining to open-mindedness to be inconclusive.

ABSTRACTOR'S ANALYSIS

It is encouraging that the authors made a sincere attempt to evaluate the curriculum change that was implemented. Too often changes are made with no attempt to determine whether the change has the desired effect. It would also appear that the variables which were examined in this evaluation are important ones. Most science educators would consider critical thinking ability, open-mindedness, and knowledge of the processes

of science to be worthwhile goals of science instruction. It is certainly appropriate to see how these variables might be affected by the introduction of the new Guide.

Unfortunately, the constructs examined represent multi-faceted skills and attitudes which are not well defined. We might say that they represent concepts of low validity. What one person describes as critical thinking, another does not; when I describe open-mindedness, I find that others disagree; there are even differences of opinion about what constitute the processes of science. This is not an argument against trying to assess such elusive qualities but it is a cautionary note concerning the way one interprets the assessment.

The readers of this report should be aware that critical thinking, open-mindedness, and science process skills are operationally defined by the instruments used in this study and that other measures of the same constructs might produce different results. The constructs are not measured in any absolute sense.

Although I have not used the RDS or the WISP, I used the WGCTA in my thesis research and became keenly aware that the instrument only measures certain aspects of critical thinking. Articles by Wallen (1963), Ennis (1958), and Rust (1962), which I reviewed in 1965 provide ample reason to accept with caution any suggestion that the WGCTA tells us all that we want to know about critical thinking. This is not to say that it tells us <u>nothing</u> that we want to know.

My point is simply that one needs to look carefully at the tests used in a study of this kind to be sure that the test is measuring what the curriculum developers thought they had put into their Guide. If it seems to be doing that, the test results are worth noting; if it does not, the results may be disregarded as irrelevant. The same suggestion, of course, goes for any reader who plans to use the result of the evaluation to make administrative decisions concerning the use of the Guide that was developed.

Apart from the inherent limitations of the design employed in this study, the methodology seems to be appropriate. I see no reason to suggest that the results are spurious. An attempt was made to check the representativeness of the sample. The matching techniques that were used to obtain the control group probably produces as much control as possible under the experimental conditions. (Random assignment is seldom possible in school settings. We usually have to make do with intact groups.)

I would have preferred to see the authors use analysis of covariance (using pretest scores as the covariate) when comparing posttest scores for the experimental and control groups. Of the three possible approaches to the comparison (simple comparison of posttest scores, comparison of gain scores, and analysis of covariance), the covariate analysis appears to be most appropriate. (For an excellent discussion of this issue, see Lord (1963).)

I would have been more comfortable with the study if the report had been more complete. Some of the things that would have helped me assesss the validity and importance of the findings are the following:

1) Sample test items which suggest the kind of skills being assessed. This would help the reader decide if he is concerned about the result obtained. It is possible to dig up the tests and examine them but that takes time and effort. Few of us will do it.

2) Additional information about the experimental conditions would help. The person who conducts a study is often aware of conditions that might have influenced the result. The reader should be told about them. This need not be an elaborate description. I am willing to believe an author who tells me that students appeared to take a test seriously, but when results suggest that this could have been a problem and there is no word from the author, I am left to wonder.

3) Tables of means (with maximum possible scores) and standard deviations should always be included in a report. There are numerous reasons that the means and standard deviations are

important. In this study a table of means would certainly help the reader interpret the ambiguous results for the RDS and WISP tests. For example, if the significant difference between experimental and control groups on the posttest of the RDS resulted from a decrease in score for the control group, it is probably because the control group didn't take the test seriously rather than because the experimental treatment was effective. The means could help one decide.

Even when there is no doubt that a reported difference is real, a look at the means can influence how one uses the result. Suppose, for example, that the significant difference in critical thinking ability that is reported in this study was a very small difference. (A difference of a fraction of a point may be statistically significant if the test used is highly reliable and the number of subjects in the study is large.) A reader may feel that even though the difference cannot be attributed to chance, it is still too small to merit the expense and effort required to adopt the innovation being reported.

Perhaps the journal reviewers are more responsible for deficiencies in reports than are authors. Having served on the Editorial Board of <u>JRST</u> for several years, I know that authors are urged to be brief and that some reviewers see no need to clutter up a report with ANOVA tables, means, standard deviations, and descriptive information that is crucial for an accurate understanding of the research reported.

Perhaps authors and reviewers alike need to be reminded that the purpose of research (and research reports) is to provide new information that others might use. If we keep this idea foremost in our minds and prepare reports so as to maximize their value to readers rather than maximize their value to writers--i.e., produce the maximum number of publications in the minimum amount of space--we would all benefit more from research that is done.

In the spirit of making research worth more, the authors could have done one thing. They could have suggested what it is about the Guide that they evaluated which seemed to produce the gain in critical thinking ability. The only people who can benefit from evaluative data which tell us that a given product is good or bad are those who are considering using the product--the Guide in this case. If, however, evaluators can provide some hint about <u>why</u> the product is good or bad, many more of us can benefit.

REFERENCES

Campbell, D. T., and J. C. Stanley. "Experimental and Quasi-Experimental Designs for Research on Teaching." In <u>Handbook of Research on Teaching</u>, N. L. Gage (ed). New York: Rand McNally. 1963.

Ennis, R. W. "An Appraisal of the Watson-Glaser Critical Thinking Appraisal." <u>Journal of Educational Research</u>, 52:155-158, 1958.

Lord, F. M. "Elementary Models for Measuring Change." In <u>Problems in Measuring Change</u>, E. W. Harris, (ed). Madison: University of Wisconsin Press. 1963.

Rust, V. I., S. Jones, and H. F. Kaiser. "A Factor Analytic Study of Critical Thinking." <u>Journal of Educational Research</u>, 55: 253-259, 1962.

Wallen, N. E., W. F. Naubrick, and Ian E. Reid. "Outcome of Curricular Modifications Designed to Foster Critical Thinking." <u>Journal of Educational Research</u>, 56:529-534, 1963.

McLeod, Richard I.; Berkheimer, Glenn D.; and Richard W. Robison. "The Development of Criterion - Validated Test Items for Four Integrated Science Processes." Journal of Research in Science Teaching, 12: 415-421, 1975.

Descriptors--Criterion Referenced Tests; *Evaluation; Science Education; Secondary Education; *Secondary School Science; *Scientific Methodology; *Tests; *Test Construction

Expanded Abstract and Analysis Prepared Especially for I.S.E. by R. L. Doran and Samuel J. Alaimo, State University of New York at Buffalo.

Purpose

The purpose of this study was to "develop group test items that would be good predictors of a student's success in actually performing the process in question in an experimental situation."

Rationale

Many elementary science curricula aim to develop process skills, but few instruments to assess these goals exist. Those that are available (e.g., the SAPA Individual Competency Measures, the BSCS Processes of Science Test and Tannenbaum's Test of Science Process), "have restricted themselves to face validity." The authors attempted to develop test items that exhibit concurrent validity with children's actual performance. Their thesis was that "face validity is insufficient for instruments of this type."

Research Design and Procedure

These researchers chose four integrated processes from the SAPA program: controlling variables, interpreting data, defining operationally, and formulating hypotheses. They developed objectively scored test items to be administered to groups with the goal of measuring these same four processes. The collection of items, called the Group Test, was subjected to the "normal procedure of writing and revising with the input of expert opinion, until the authors were confident that the retained items exhibited

face validity." These 79 items (multiple choice or numerical fill-in) were supplemented with 42 slides and a taped oral script.

A group of 56 seventh graders who had the SAPA instruction as sixth graders was administered the Group Test and individually assessed for their performance on 54 tasks from the SAPA Individual Competency Measures of the four processes of interest. A comparison of the items or tasks to the respective processes is listed in Table 1.

Distribution of Items from the Group Test and ICM tasks by Process Areas

	ICM	Group Test
Controlling Variables	12	18
Interpreting Data	15	25
Formulating Hypotheses	14	18
Defining Operationally	13	18
	54	79

Individual testing for each student on the ICM tasks required two 50-minute periods. Following the 2-1/2 month ICM testing phase, the entire sample was administered the Group Test.

Indices of the difficulty and discrimination for each item were computed on the Group Test. The external criterion for the discrimination calculations was the student score on the ICM tasks. After deleting items with discrimination indices below +0.20, the investigator calculated bivariate correlation coefficients between and among the process subtests from the Group Test and ICM tasks assessing similar process.

Findings

The experimenters found that the mean indices of discrimination for each process subtest of the Group Test ranged from +.17 to +.33 when 22 items

(of the 79-item test) that had discrimination indices below +.20 were deleted, the mean discrimination indices for the four process subtests rose dramatically to .28, .35, .39 and .44, as expected. The level of difficulty of the items in these subtests before and after deletion was "relatively unchanged and very close to the desired level of 50 percent."

The correlation coefficients between the similar subtests on the Group Test and ICM were all positive and statistically significant at the 0.0001 level. The coefficients ranged from +0.535 to +0.705. Coefficients between unlike subtests of the two tests were also positively related, and all significant at the 0.01 level. These ranged from +0.414 to +0.703. Similarly, correlation coefficients among the four subtests of the Group Test and among the four ICM subtests were all positive and significant, ranging between +0.430 and +0.601 for the ICM and +0.561 to +0.786 for the Group Test.

Interpretations

The authors concluded that it is "possible to produce objective items on a test which exhibit high correlation with a student's ability to actually perform the higher mental processes as represented by performance on the integrated processes."

"The likelihood that the integrated processes are, indeed, not unique is strongly suggested by the intercorrelations within each test." The authors speculated that success on the integrated processes may be predicted by a subset of basic processes, level of cognitive development (à la Piaget), reading ability or the student's mental ability. They suggested that one of these as yet unidentified variables may be influencing these strong interrelationships.

The authors subsequently were to choose some of these items and form a group process test, complete with normative data. They felt this test could be used by researchers who wish to assess the process skills of students in various science programs. They hoped their work would stimulate research to identify factors which influence the development of science process skills and to explore the generalizability of these skills to other content areas.

ABSTRACTOR'S ANALYSIS

The problem investigated by this project is significant to science education and is stated clearly and succinctly. The review of related literature was very limited. Several attempts to assess student ability to perform SAPA process skills have been reported by Beard (1971) and Wallace (1974). Even more specifically, Walbesser and Carter (1970) investigated the effect of test format (Individual and Group) on student's performance of process skills. Analysis of these studies may have been of some help to the conduct of the study or in the preparation of the report.

The small sample size limits the generalizability of the findings of the study. The students were only described as being seventh graders who had studied SAPA in the sixth grade. It's not clear if these students had SAPA instruction during any or all of their K-5 science experiences. The ability level of the students similarly was not specified. Nor was there a brief comment as to the type of school or community(ies) from which these students were selected. Since most schools have considerable information on reading and mental abilities of students, this information could have been easily gathered from appropriate school folders.

The choice of tasks from the ICM seems logical and appropriate. It can be assumed some of the investigators administered and scored the ICM tasks, but this was not described in the report. It appears that the Group Test was administered after all students had been individually administered the ICM tasks. The potential instructional effect of this lengthy assessment procedure was not discussed. The range and central tendency of student scores on these tasks was not included in the report. It would have been helpful to append an illustrative example of ICM tasks and items from the Group Test. Before one computes correlation coefficients, it would seem appropriate to describe some characteristics of the data involved.

The analysis was relevant and appropriate, considering the sample size. No hypotheses were stated with respect to difficulty levels, discrimination indices or correlation coefficients. With a non-SAPA control group, one could compare means on specific processes, etc. The minimum level

for "acceptable" index of discrimination of +0.20 was arbitrary but appropriate. However, the "desired level of 50 percent" difficulty for items seems consistent with norm-referenced systems but not with the expressed intent to produce criterion referenced items.

The use of an external criterion, performance on the ICM tasks, is an appropriate procedure. The size of the ICM subtests was relatively small, from 12 to 15 tasks, so one wonders about the reliability of these scores. Reliability estimates of subtests and total tests were not reported for either of the two instruments. It would be interesting to know if there were any similarities among the deleted items (or among the remaining ones), in terms of content assessed, reading demands, perceptual demands, mathematical complexity, etc. If there was a larger sample, one could explore commonality among items via a factor analysis and similar techniques.

The 57 items retained, called "successful" by the authors, were the basis of calculation of correlation among the process subtests of the Group Test and the subtests of the ICM. The results indicating non-uniqueness of the process as assessed by both the ICM and Group Test were disappointing.

More sophisticated means of analyzing the data may not have been appropriate with this sample, but it seems logical to consider using multiple regression (MR) or canonical correlation techniques. With MR one could determine the relative contribution from each of the ICM process areas to the four processes as assessed in the Group Test. This technique would also be most useful when data from additional variables are available to quantitatively compare the relative contributions. Canonical correlations are appropriate when one has a set of predictor variables and a set of criterion variables, which is true in this study. Outcomes from such an analysis would enable one to determine variance common to the two sets of variables as well as ways in which the left and right sets of variables are individually structured. One could obtain information assto the individual contribution of each variable from each set to the interrelationships among the two sets.

The assessment of process skills is an important area for research by science educators. There are many uses to which research findings and measurement tools in this area can be applied. It is hoped that this

research is viewed as a first step of a multi-staged domain. Exploration of the relative performance of students from various programs and transfer to other content areas were suggested by the authors. The contributions of student achievement, psychomotor skills, science knowledge, or "interest in sciences" to student process ability need to be explored. One could investigate the relative effects of various instructional programs on students in general, or students with specific aptitudes. The effect of process ability or success in concurrent and future science courses, post-secondary plans, etc. could be explored when some prior studies have been completed. It is hoped that these investigators will continue to wrestle with these research problems.

REFERENCES

Beard, Jean. "The Development of Group Achievement Tests for Two Basic Processes of AAAS Science--A Process Approach." Journal of Research in Science Teaching, 8(2):179-185, 1971.

Walbesser, Henry H. and Heather L. Carter. "The Effect on Test Results of Changes in Task and Response Format Required by Altering the Test Administration from an Individual to a Group Form." Journal of Research in Science Teaching, 7(1):1-8, 1970.

Wallace, Charles. "ERIE Science Processes Test" in Unpublished Evaluation in Science Education: A Handbook, by Victor J. Mayer. ERIC Science, Mathematics and Environmental Education Clearinghouse, The Ohio State University, Columbus, 1974.

INDIVIDUAL STUDIES

Thorsland, M. N. "Formative Evaluation in the Development of an Audio-Tutorial Physics Course." Science Education, 59(3):305-312, 1975.
Descriptors--*Autoinstructional Methods; *College Science; Educational Research; Higher Education; *Individualized Instruction; *Instruction; *Physics; Science Education

Expanded Abstract and Analysis Prepared Especially for I.S.E. by J. Dudley Herron, Purdue University.

Purpose

The study involves formative evaluation of an Audio-Tutorial university physics course. This evaluation focused on teacher-student and student-student interaction during the course and attitudes of students toward the laboratory and toward the course in general.

The following hypotheses were tested:

H-1 A higher percentage of A-T students than non A-T students will respond favorably to statements pertaining to attitudes toward the course in general.

H-2 A higher percentage of A-T students than non A-T students will respond favorably to statements pertaining to attitudes toward the laboratory.

H-3 A higher percentage of A-T students than non A-T students will respond favorably to statements pertaining to the availability of personal help and personal contact in the course.

H-4 There will be no significant differences in achievement between the A-T group and the non A-T group.

Rationale

The rationale for the study is course improvement. Conventional classes were assumed deficient in their lack of syncronization between laboratory and lecture and in their lack of student involvement during large lectures. Audio-Tutorial lessons were instituted to overcome these weaknesses and the study was initiated to determine the effectiveness of the new program.

Research Design and Procedure

Seventy students were randomly selected from a class of 420 and were instructed using A-T methods for ten weeks. The rest of the class (N=350) studied in a conventional lecture, recitation, laboratory format.

Mean SATV and SATM scores for the treatment and control groups were compared and found to be virtually identical. In addition, a questionnaire was administered which asked about college major, courses taken in high school, grades obtained, and grade aspirations in physics. No differences between the treatment and control group were found on any of these characteristics.

A 15-item questionnaire was administered to both treatment and control groups after seven weeks of instruction and the results were analyzed to test hypotheses 1 through 3. Means on the three examinations given during the course were compared to test hypotheses 4. The chi-square statistic was used to test for significant differences in response to each item on the 15-item questionnaire and ANOVA was used to test differences in means on the three examinations.

Findings

There were five items on the questionnaire which pertained to general attitudes toward the course. On four of the five items, the response of the A-T students indicated a more favorable attitude than that of the conventionally taught students ($\alpha \leq .01$) Two items provided information about attitudes toward the laboratory and both produced responses which indicated more favorable attitude on the part of the A-T students. ($\alpha < .001$) Three items dealt with personal contact. All three indicated that A-T students were more convinced than conventional students that they had sufficient opportunity for contact with instructors and other students ($\alpha \leq .01$). Thus, H-1, H-2, and H-3 were all accepted.

The results on the achievement tests were not as encouraging. There were no statistically significant differences in means for any of the three examinations. However, the author pointed out that there appeared to be slight differences in favor of the conventionally taught students on each of the three exams (7 percent higher on Exam I, 6 percent higher on Exam II, 4 percent higher on Exam III). Even though none of these differences was large enough to rule out the possibility that it occurred by chance alone (α = .05), the author considered it to be of some practical significance and attributed the difference to an emphasis on problem solving in the conventionally taught course.

Interpretations

The author concludes that "results of the attitudinal testing indicated that significantly more A-T than non A-T students responded favorably to questionnaire items related to attitudes toward the course in general, toward the laboratory, and toward personal contact in the course." . . . "overall, the A-T group's achievement was slightly below the non A-T group on course exams. The differences were not statistically significant. . ."

ABSTRACTOR'S ANALYSIS

Formative evaluation is done for the benefit of those teaching the course. According to the author, the exploration of the audio-tutorial format continued after this study ended and "the course, as presently conducted, retains a basically audio-tutorial format with self-pacing and mastery-learning aspects and serves approximately 600-700 students each semester." The author's statement implies that the course was revised on the basis of the evaluation reported. If so, this is as it should be. Formative evaluation is done for pragmatic reasons. The data are to be used for revision.

Data obtained from formative evaluation are usually of limited value to "outsiders." Conditions vary from place to place and classroom practice

must be influenced by classroom space, faculty load, scheduling patterns, budget, and other logistical considerations. It is difficult to report conditions under which a course was taught in enough detail for a person at a distant campus to decide the probability of obtaining similar results from the same practice.

This is not to say that studies of this kind are of no value. If enough are reported, general patterns begin to emerge.

Based on studies that I have seen, it is probably that the more favorable attitudes expressed by A-T students in this study are representative of attitudes in general. This, of course, will vary from one situation to another. A bad A-T program will not elicit more favorable reviews than a stimulating lecture course but when both are well run, the A-T is likely to get the higher marks.

The finding of "no significant difference" in achievement reported in this study is also consistent with results reported for other evaluations of A-T instruction.

In a recent article, Kulik and Jaksa[1] report finding 24 studies of A-T instruction which they considered "experimentally adequate." Of these 24 studies, 9 reported significantly higher final exam scores in A-T classes, 2 reported superior performance in the conventional classes, and 13 reported no significant difference in achievement. Thus, there is strong evidence that, on average, students achieve no less in A-T classes than in conventional classes and there is weak evidence that they may achieve more. It should be pointed out, however, that Kulik and Jaksa do not report the subject matter being taught in the 24 studies that they reviewed. It is quite possible that A-T instruction works better for some subjects than for others. Work done in an engineering course which stressed problem solving skills has suggested that A-T may not work as well in courses that stress analytical thinking as in courses that stress information acquisition. This reviewer's impressions gained from reviewing research on A-T over the past several years is consistent with this.

In addition to evidence that A-T courses generally receive favorable comments from students and that achievement in A-T courses is equal to or slightly greater than achievement in conventionally taught courses, there is some evidence that time required for A-T courses is less than in conventional courses. The data available on this point are meager since most research on A-T does not report information about time. However, Kulik and Jaksa (1977) report that four of the five studies that did provide this information reported substantial reductions in the amount of time needed for students to complete the course work. The study being reviewed here does not report information on this point.

This reviewer has repeatedly suggested that evaluation studies which ask, "Is method A better than method B?" are of limited value. If students in A-T courses have more favorable attitudes, achieve more, or take less time to accomplish the course objectives, we need to know what it is about the A-T approach that leads to these benefits. Knowing that, we are in a better position to tell instructors at other institutions what they need to do to obtain similar results. Unfortunately, most evaluation studies fail to give us this information and we are left to guess.

It seems likely that the greater flexibility of the A-T format may account for the more favorable attitudes found in this study. As one colleague put it, "Students like options." The A-T student certainly has more options than does the student in the conventional class. Generally he has freedom to schedule his own time, to sequence learning activities and perhaps even to decide which units are to be completed. This is seldom the case in conventional classes.

It seems that the integration of laboratory and didactic material in the A-T format also leads to more favorable attitudes. In the study under review it was reported that almost 80 percent of students in non-A-T classes indicated that most lab work was not very informative considering the time spent on it. In contrast, only 40 percent of the A-T students felt this way. In addition, 55 percent of the non A-T students in the study indicated that they had difficulty relating the lab work to the material from the rest of the course while only 26

percent of the A-T students expressed this feeling. Such differences could have a marked effect on the attitude of a student toward a course.

An interesting bit of information provided by this formative evaluation was not discussed by the author—undoubtedly because it did not indicate a difference between A-T and conventional instruction. It was found that almost 70 percent of students in both A-T and non A-T classes indicated that they would like more opportunity to "mess around with physics apparatus and engage in non-quantitative experimentation." Is what these students are telling us important? Perhaps our formative evaluation should be asking whether our rush to develop quantitative skills so important in the physical sciences has robbed students of opportunities to make the qualitative observations needed to understand the point of quantitative work. As science educators we seem to give far more attention to instructional techniques than to content of instruction. One hears many heated debates on the relative merits of descriptive and theoretical topics in a beginning course but there is virtually no hard (I would settle for slightly gelled!) data bearing on the debate. Is it possible to obtain? I think so. It would certainly be worth some exploratory research to find out.

REFERENCE

Kulik, J. A. and P. Jaksa. "PSI and Other Educational Technologies in College Teaching." Educational Technology, 17(9):12-19, 1977.

Voelker, Alan M. "Elementary School Children's Attainment of the Concepts of Physical and Chemical Change--A Replication." Journal of Research in Science Teaching 12(1):5-14, 1975.
Descriptors--Chemistry; *Concept Formation; *Educational Research; Elementary Education; Elementary School Science; *Instruction; *Maturation; Science Education; Teaching Concepts

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Donald E. Riechard, Emory University.

Purpose

The purpose was to determine the relative effectiveness of two methods of instruction in teaching the concepts of physical and chemical change. An associated question was to determine whether maturity of the children as represented by grade level was related to the children's level of understanding of the selected concepts.

Rationale

There are two parts to the rationale developed for this study. The first deals with the need for research evidence to support judgments about the placement of science concepts in the schools and how the related instruction should be organized. The second part makes the case for replication of studies in science education. This study is a replication of one previously completed by the author (Voelker, 1968).

Research Design and Procedure

The general research design involved application of two different treatments to two different groups of subjects from each of three grade levels. There were no pretests. Subjects were post tested after treatment.

The population for this study consisted of all the children in grades four, five, and six of a single elementary school in a large midwestern community. Each grade level had four separate classes. Two classes from

each grade level were randomly selected to participate in the study. The instructional treatments (treatment-one or treatment-two) were randomly assigned to the two classes. Thus, six classes were involved in the study. Following instruction, ten children from each of the six classes were randomly selected for testing.

The primary difference in the two treatments was in the roles played by teacher and students. In treatment one (T1), the responsibility for formulating (discovering) the basic scientific generalization was on the student. In treatment two (T2), the teacher formulated and stated the generalization.

Experimental procedures and testing were conducted over four consecutive weeks. The author describes the testing as a "semi-clinical approach." Individual children were asked to: 1) answer three verbal questions about physical and chemical change; 2) classify demonstrated and described phenomena as examples of either physical or chemical change; and, 3) justify the classification of each phenomenon.

Analyses of the data were summarized in six tables. Analysis of variance (ANOVA) was the primary statistic used. A discussion of results is keyed to the information presented in the tables.

Findings

Answers to Verbal Questions. At the fourth grade level the mean scores of the T1 group were greater than those of the T2 group for each of the three questions. At the fifth and sixth grade levels, the mean scores of the T2 group are greater than or equal to the mean scores of the T1 group for each of the three questions. However, none of the means was significantly different (alpha = 0.05) for either main effects or interaction effects.

Classifying Demonstrated and Described Phenomena. Fourth and sixth grade T1 group mean scores on demonstrated phenomena exceeded those for the respective grades' T2 group. The mean scores for the T1 group were greater than or equal to those of the T2 group at all grade levels when classifying described phenomena. Significant differences (alpha = 0.05) for the grade level effect were found for both described and demonstrated phenomena. For demonstrated phenomena, the grade six mean was significantly greater than the grade five mean whereas the grade six mean was significantly greater than both the fourth and fifth grade means on described phenomena (Newman-Keuls, p = 0.05).

An item analysis revealed that eight of the fourteen demonstrated phenomena were correctly classified by 50 percent or more of the children receiving T1 or T2 treatments. None of the six described phenomena was classified correctly 50 percent or more of the time by children receiving either T1 or T2.

Justifying the Classification of Phenomena. At both the fifth and sixth grade levels, the mean scores for the T2 group were greater than the mean scores for the T1 group. There was a significant difference (ANOVA; p = .04) between the means for the grade level factor only. Although students had the opportunity to ask questions as a means of gathering more information about the phenomena, the percent of instances where questions were asked was less than 1 percent.

Minimum Success Level. Minimum success levels were set for each of the three parts of the test (questions, classification, justification). Only the fourth grade T1 group met the minimum criterion for answering the questions. The fourth grade T2 group met the minimum criterion for classifying the two types of phenomena. Both sixth grade T1 and T2 groups met the minimum criterion for classifying demonstrated and described phenomena and supporting correct classification with acceptable responses.

Correlations. Correlations between children's scores on correct classifications of phenomena and support of correct classifications with acceptable reasons were significant ($p = 0.05$) for four of the six grade level-treatment groups (G4-T1, G5-T2, G6-T1, G6-T2). Correlations between test scores and standard achievement scores (reading vocabulary and comprehension, arithmetic computation, spelling, and IQ were significant ($p = 0.05$) for the grade 6-T1 group only.

Interpretations

The author states that this replication study has added credibility to the following conclusions from the original investigation:

1. It does not appear appropriate to attempt to teach the concepts of physical and chemical change prior to grade six if instruction is to be in large group, teacher directed format, and the expectation is that children meet minimum criteria for all objectives set forth.

2. A major inhibitor to concept formation appears to be maturation as a function of grade level.

3. Children are more able to reveal their understanding of the concepts of physical change and chemical change through classifying phenomena rather than formulating definitions or applying concept definitions.

4. Children can be expected to generate some knowledge for themselves.

5. Children are not learning how to ask questions that will help them learn.

6. There is a relationship between the ability to classify phenomena correctly and being able to explain why the respective classifications were made.

7. The relationship between success in developing these concepts and achievement on standard tests appears to be more a function of the classroom in which the student is found rather than a general factor across age level or treatment group.

ABSTRACTOR'S ANALYSIS

Relationship to Other Studies. This investigation is related to a large cluster of studies on concept learning. The original investigation (Voelker, 1968) whose replication is reported here was conducted through the Wisconsin Research and Development Center for Cognitive Learning at Madison. Numerous science concept studies came out of that R & D Center at about the same time. Among them were those by Carey (1968), Helgeson (1968), Pella and Ziegler (1967), Stauss (1968), and Triezenberg (1968).

It is not feasible here to produce a complete listing of investigations related to this study and the general area of concept learning in science. It is noteworthy, however, that published research expressing concern for children's cognitive development in science can be traced back many years. Hall's study, for example, dates into the last century (1891). The work of Piaget and his associates, begun in the 1920's, surely is of much significance. Examples of some of the more recent studies related to science concept learning are those reported by Anderson (1965), Haney (1965), Howe (1974), Lawson & Wollman (1976), Raven (1974), and Riechard (1973).

Because of the voluminous nature of the concept learning literature, the reader is urged to enlist the services of ERIC/SMEAC in the search for specific research related to his or her own study. Other sources (some available through ERIC/SMEAC) are the various concept learning bibliographies such as those by Klausmeier, et al. (1969) and Voelker (1973).

Contributions of the Study. The results of this study are generally consistent with that which is known about cognitive development. Further,

taken with the original investigation, there is good evidence to suggest the appropriate grade placement for the concepts of physical and chemical change. It is this abstractor's belief, however, that a major contribution of the investigation lies in the fact that it is a *replication*. The author makes a sound case for replication studies and it is likely that most science educators agree with him. However, a cursory review of the literature shows that very few replication studies ever appear (probably less than five percent of the total research published). Investigators and those responsible for directing research (e.g., graduate student research, etc.) should consider the need and potential contribution of replication studies. It seems that replication of selected works could make a far more substantive contribution than the continuation of one-shot, bits-and-pieces research.

Current State of Research in the Field. As suggested above, much research has been done on concept learning. A good bit of it has been done in science education. Four basic concerns can be identified relative to that research. First, different studies are seldom tied together as they are planned and produced or after the results are published. Some positive signs can be identified, however. The Wisconsin R & D Center, as mentioned above, made substantial progress in planning and producing related research in science concept learning. More efforts of this nature are needed. And this ERIC publication, Investigations in Science Education, is a welcome effort toward organizing and examining "clusters" of related research that have appeared in the literature.

Second, many of the studies have been one-shot research. Most have not been replicated; some could not be replicated. Stuides need to be re replicable and replicated!

Three, much of the research has been done with children seven years of age or older. Given the psychological research which suggests the importance of intellectual growth prior to age seven, there is need to study concept development in the early years.

And finally, even considering the three concerns above, more is known about concept learning in science than is practiced in the classroom, Too many teachers, for example, still try to teach science concepts by the single method of requiring students to read the chapter and write out the answers to questions at the end. Thus, there is need to decrease the lag time between research and practice.

Comments on the Study. The research design for this study is adequate and, in general, the written report is well done. Language is precise and the material presented is organized into sections which flow smoothly from one to the next. The data appear to be analyzed thoroughly and appropriately.

As with all reports, however, some questions and concerns can be identified. Neither the number of students making up the total population nor the number of subjects in each treatment group is stated, for example. How large was each class and was each class instructed as a total unit? The number in each instructional group is of major importance given the reference to group size in interpretation number one, above.

The liberal use of tables provides an excellent aid in the examination of results. Tables IV and V are not completely clear at first reading, however, and require more careful study than the other tables. Further, the discussion of *item analysis* implies that Table V deals with both demonstrated and described phenomena; in fact, it deals only with demonstrated phenomena. It is also noted that the relationships among the grade-level scores on the described phenomena are missing from the footnote in Table II.

Finally, the author appropriately related his conclusions (see section on *interpretations*, above) to the original investigation. This should be done in replication studies. While the results of this study generally support the conclusions from the original investigation, the replication data are not without incongruences. For example, the fourth grade groups

frequently outperformed the fifth graders and even the sixth grade groups on a few measures. Why? The question cannot be answered here. One possible answer is sampling bias. In this study, the sample was selected from one school only, where students had been assigned to classes prior to the investigation. Criteria for selection of the school and assignment of students to classes were not defined. There were no pretests to determine student performance prior to instruction. The author made good use of randomization on those things over which he had control. The study would have been more generalizable with less chance of sampling bias, however, had the population been larger and more diverse, if all assignments had been random, and if pretests had been used. The reader should recognize, however, that restraints are often imposed on behavioral research. The researcher frequently has to work with "what is" rather than "what ought to be." While some questions about sampling bias in this investigation can be raised, they do not detract greatly from the overall merits of the study.

REFERENCES

Anderson, R. D. "Children's Ability to Formulate Mental Models to Explain Natural Phenomena." Journal of Research in Science Teaching, 3(4): 326-332, 1965.

Carey, R. L. Relationship Between Levels of Maturity and Levels of Understanding of Selected Concepts of the Particle Nature of Matter. Technical Report No. 36. Madison: Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin, 1968.

Hall, G. S. "The Contents of Children's Minds on Entering School." Pedagogical Seminary, 1:137-173, 1891.

Haney, R. E. "The Development of a Non-verbal Test of Children's Concepts of Animals." Journal of Research in Science Teaching, 3(3):198-203, 1965.

Helgeson, S. L. The Relationships Between Concepts of Force Attained and Maturity as Indicated by Grade Levels. Technical Report No. 43. Madison: Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin, 1968.

Howe, A. C. "The Acquisition and Growth of the Concept of Conservation of Liquid." Journal of Research in Science Teaching, 11(2):105-112, 1974.

Klausmeier, H.; B. Sterrett; D. Frayer; S. Lewis; V. Lee; and J. Bavry. Concept Learning: A Bibliography, 1950-1967. Technical Report No. 82. Madison: Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin, 1969.

Lawson, A. E. and W. T. Wollman. "Encouraging the Transition from Concrete to Formal Cognitive Functioning--An Experiment." Journal of Research in Science Teaching, 13(5):413-430, 1976.

Pella, M. O. and R. E. Ziegler. The Use of Static and Dynamic Mechanical Models in Teaching Aspects of the Theoretical Concept The Particle Nature of Matter. Technical Report No. 20. Madison: Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin, 1967.

Raven, R. J. "Programming Piaget's Logical Operations for Science Inquiry and Concept Attainment." Journal of Research in Science Teaching, 11(3):251-261, 1974.

Riechard, D. E. "Life-Science Concept Development Among Beginning Kindergarten Children from Three Different Community Settings." Journal of Research in Science Teaching, 10(1):39-50, 1973.

Stauss, N. G. An Investigation into the Relationship Between Concept Attainment and Level of Maturity. Technical Report No. 40. Madison: Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin, 1968.

Triezenberg, H. J. The Use of An Advance Organizer in Teaching Selected Concepts of Ecological Systems. Technical Report No. 42. Madison: Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin, 1968.

Voelker, A. M. The Relative Effectiveness of Two Methods of Instruction in Teaching the Classificational Concepts of Physical and Chemical Change to Elementary School Children. Technical Report No. 39. Madison: Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin, 1968.

Voelker, A. M. Research on Concept Learning: An Analysis. Item No. 133S. Columbus: Center for Science and Mathematics Education, The Ohio State University, 1973.

64

Pizzini, Edward L. "The Effects of An Exploratory Teaching Experience on Selected Educational Concepts." Science Education, 59(2):207-210, 1975.

Descriptors--*Attitudes; Educational Research; Higher Education; Individualized Learning; *Preservice Education; Science Education; *Science Teachers; *Teacher Education; *Teaching Procedures

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Victor J. Mayer, The Ohio State University.

Purpose

The purpose of the study was to determine the effects of a teaching experience upon undergraduate students' attitudes toward teaching and certain instructional practices. The instructional objectives evaluated were not listed, however they can be inferred from the "basic rationale" for the experience:

1) to allow the students enrolled to make a rational decision pertaining to the feasibility of teaching as a career, and

2) to provide first-hand experience of teaching to insure maximum gain from advanced, professional education coursework to be taken.

Rationale

Not provided.

Research Design and Procedure

The design used was the one group pretest and posttest design. The instrument used was a semantic differential to assess changes in attitudes. It had eleven concepts and a scale consisting of ten adjective pairs each with seven semantic spaces. A twelfth variable was determined by summing the scores of the eleven concepts.

The experience evaluated consisted of a one-semester participation in the daily activities of a public school. It enrolled 21 sophomores at the University of Iowa.

Findings

A positive change in the pretest and posttest scores, significant at the .05 level of confidence, was found for 11 of the 12 variables.

TABLE I
Pretest and Posttest Results of Attitude Toward Selected Educational Concepts as Measured By The Semantic Differential for Exploratory Teaching Participants

	Pretest Mean	Posttest Mean	t
Individualized Learning	22.67	11.48	6.39*
Being a Science Teacher	25.38	15.38	5.15*
Teaching Secondary Students	25.24	14.81	6.45*
Interaction	21.05	7.62	7.25*
Content-Oriented Approach	33.57	26 86	3.63*
Classroom Mangement	35.43	30.10	2.84*
Science Teaching Materials	25.05	14.00	5.52*
Teaching Elementary Students	25.05	12.43	8.09*
Process-Oriented Approach	22.76	11.67	6.69*
Importance of Discipline.	31.81	28.57	1.76
Early Exploratory Teaching	23.95	15.57	3.11*
Total Composite Score	293.33	188.76	11.54*

Critical Region = 1 1 > 2 08
*Significant at the 0.05 level N = 21, df = 20

Interpretations

The exploratory teaching experience resulted in a positive change in attitudes.

The following are additional interpretations made by the author:

> The changes are indicative of the development of a more sensitive, confident, humanistic, and laboratory-oriented prospective teacher. Furthermore, the objectives of the program were apparently fulfilled as the data would indicate that the participants have developed insights into teaching which should allow them to make more rational decision (sic) as to whether teaching is a potential career.

ABSTRACTOR'S ANALYSIS

No rationale is provided for the use of an exploratory teaching experience; no summary of appropriate studies. The reader is left to his own devices to try to determine what led the author to providing such an experience for his students. There is substantial literature in science education bearing upon this question.

No rationale is provided as to why an exploratory teaching experience should affect student attitudes toward teaching. Again there is substantial literature. I would hope that at least those studies using a semantic differential, such as the one by James (1970) would be cited. Only one, however, relating to preservice training is cited in the introduction. No attempt is made to relate the results of this study to those of others having a similar purpose.

No discussion is provided concerning the methods used in the development of the instrument, measures used to insure validity, nor estimates of reliability. No discussion of the dimensions commonly comprising semantic differentials (evaluative, potency, activity and understanding) and found to be relatively stable is provided as an underpinning to the development of the instrument. It would seem that five of the adjective pairs are evaluative, two are understanding, one is potency and two do not seem to fit any of the four dimensions. What was the rationale for this distribution? Do the four dimensions show different patterns of response? Was a factor analysis attempted? How were the concepts selected? The adjective pairs? They relate to the validity of the instrument. The author does state that the concepts selected were determined by field testing. However, no information is given on the procedure used.

The interpretations offered by the author are not possible with the design used. There is a large body of literature indicating very weak, if any, relationship between attitudes and behavior. Yet the author concludes that changes in attitudes documented in the study "are indicative

of the development of a more sensitive, confident, humanistic, and laboratory-oriented prospective teacher"....implying that his students will behave in those ways. The results of the study may demonstrate the formation of attitudes that would favor the development of sensitive, confident, humanistic and laboratory-oriented prospective teachers, however, I doubt that they are indicative of such development.

The article is written in a clear, succinct style. It is easily understandable. However, in addition, to the problems previously addressed, however, there is some inappropriate use of terminology. This is not an experimental study. At best it is an evaluation of a pilot program. Yet a description of the program is entitled "experimental treatment." The lead sentence under "Analysis and Findings" states that "the null hypothesis stated that there was no...etc." The null hypothesis is never rejected. Instead, it is later stated that a significant difference was found for eleven of the twelve variables. One wonders why the term "null hypothesis" is ever introduced. Perhaps for the same reason that many investigators use null hypotheses as research hypotheses in reporting their studies; they seem to lend a certain amount of sophistication. In a descriptive study such as this they are superfluous as the author seems to concede since it is not brought up again after its introduction. Null hypotheses are statistical devices used in the analysis of data. Although not used in this study as research hypotheses, they have been so used in many studies appearing in the science education literature. Such uses are inappropriate and misleading.

REFERENCE

James, Helen Hope. "Effects of Three Supervisory Method Upon the Development of a Teaching Strategy Among Science Student Teachers". Unpublished Doctoral Dissertation, University of Colorado, 1970.